Turkish Journal of Computer and Mathematics Education Vol.12 No.10 (2021), 2677-2683

Research Article

A Relative Evaluation Of Specificity And Sensitivity In A Diseased Plant Leaf With Hybrid Sift And K-Means Fuzzy Logic Svm Algorithm

M.Sowmya[1] Dr.Bojan Subramani[2]

[1] *PhD Research Scholar,Department of Computer Science,Shri Nehru Maha Vidhyala College of arts and sciences,Coimbatore,Tamil Nadu, India.*
[2]*Principal & Associate Professor,Department of Computer Science, Shri Nehru Maha Vidhyala College of arts and sciences,Coimbatore,Tamil Nadu, India.*
phdsowmya89@gmail.com [1], drbsubramani@gmail.com [2]

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: - Early location of the plant infections is basic to maintain a strategic distance from misfortunes in the yield and nature of the farming item. The investigations of the plant illnesses have been generally explored to recognize irregularity in plant development utilizing outwardly discernible examples on the plant. Plant observing and illness location is expected to guarantee maintainability in agriculture. Be that as it may, it is generally hard to screen the plant illnesses physically as they require an ongoing and accurate location. Picture handling is usually utilized for the discovery of plant sicknesses which included picture securing, pre-preparing, division, include extraction, and characterization. In this research paper, A hybrid-based image processing algorithm is proposed to detect diseases on six datasets with plants using its leaves images. The proposed solution focuses on using hybrid k means fuzzy logic SVM and hybrid shift to classify these images. Different parameters are used to compute different algorithms. The results are classified into Sensitivity and specificity.

Keywords:- Fuzzy, SVM, SHIFT,K Means, Sensitivity, Specificity

1. INTRODUCTION

An expansion in the worldwide populace had brought about high requests for food and more concentrated cultivating rehearses by farmers. farmers are currently needed to deliver more food per section of land because of restricted farming areas. Combined with plant sicknesses happening in plants, both amount and quality yield of the food creation is in danger With early right analysis and therapy, this can be diminished or influenced zone can be reestablished back to its better condition. Specialists, for example, plant pathologists can decide the legitimate analysis and conceivable treatment the plant may have in various ways. One of the basic strategies is to screen the shade of the plant leaves. Various illnesses would bring about various shapes and shade of the influenced spots on the leaves. By utilizing PC picture preparing and machine language calculation, these influenced spots can be resolved and handled to characterize accurately the plant infection and conceivable treatment.

2. METHODOLOGY

2.1 HYBRID ALGORITHM

The fundamental characterization use for arrangement is a Support Vector Machine (SVM). SVM is one of the managed AI used to characterize informational indexes into classes. SVM discovers separators with the most extreme edge to improve the presentation of the classifier. Bit work utilized in SVM is characterized as the numerical equation use to change indivisible informational collection from the information plot to a lower or higher dimensional space that outcomes in distinct outcomes. 100 sectioned pictures are utilized to prepare the calculation with 5-crease cross approval. 5-overlay cross approval is the way toward partitioning the informational collection arbitrarily into 5 gatherings for example there are 20 arbitrary informational collections in each gathering. The preparation is then rehashed multiple times with each time 1 of the gathering is used to approve the precision of the characterization. The normal of this correctness is determined to give better exactness to the arrangement.

2.1.1 For quadratic SVM, a polynomial piece work with the polynomial request, q of 2 is utilized with computerized part scale also.

$$G(x_j + x_k) = (1 + x_j{}' x_k)^q \quad (1)$$

2.1.2 Support Vector Machine

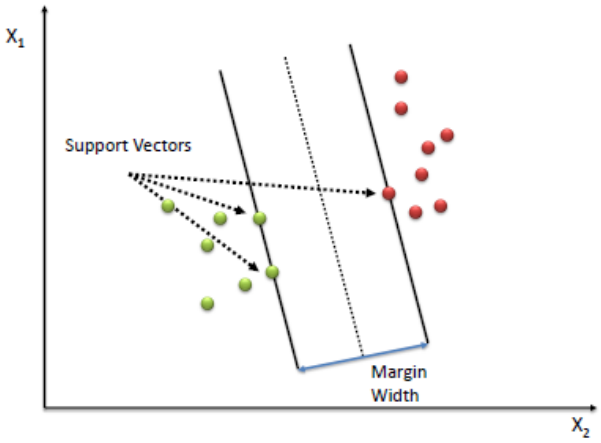

Fig 1: Separating hyperplane created by SVM.

Support Vector Machine or SVM is perhaps the most broadly utilized regulated AI. For managed learning, to test the calculation, preparation information with named classes is needed to prepare the calculation. SVM is very comparative with discriminant, however, what makes SVM stands apart is a direct result of SVMs builds the greatest edge separator that outcomes in better speculation when contrasted with discriminant classifier. For a straight distinct informational index, a direct isolating hyperplane is made. Typically, the collection of information that is not directly divisible in the first plotted dimensional space is distinct in higher-dimensional space or lower-dimensional space. A part stunt is utilized to change the first-dimensional space to a higher-dimensional space utilizing a numerical recipe called piece work. Accordingly, SVMs are known as a nonparametric technique. Even though enormous datasets are utilized for preparing the calculation, just certain highlights or informational collections are stores in the prepared calculation. For a 2-dimensional space information plot, Fig shows the essential thought on how to uphold vector machine functions.

2.1.3 SIFT

Scale-space extreme location: In request to execute the main stage, it was critical to distinguish highlights at various scales. Along these lines, one great decision was Gaussian [1] kernel (G) as shown in Equation 2:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2)$$

$$D(x, y, \sigma) = \left(G(x, y, k^{\sigma})\right) * I(x, y) = L(x, y, \sigma) - L(x, y, \sigma) \quad (3)$$

Here equation 2 addresses the distinction of Gaussian (DOG), which is viewed as a respectable guess to the scale-standardized Palladian of Gaussianσ2_2G. From that point forward, an octave of scale-space was fabricated It contains pictures with various scales as demonstrated in Figure. To discover nearby maxima and minima of DOG, a pressure activity was applied between each example point and its neighbors in the top and base size of that point.

(i) LOCALIZATION:

In this stage, there are fewer focuses to adapt to instead of utilization all picture pixels. Be that as it may, it is as yet countless key points. They can be diminished by disposing of key points [2] with low differentiation since they are touchy to the clamor. This was finished by taking Taylor extension of scale-space work as following

$$D(x) = D + \frac{\partial \mathrm{D}^{\mathrm{T}}}{\partial \mathrm{x}} * \mathrm{x} + \frac{1}{2}\mathrm{x}^{\mathrm{T}} * \frac{\partial^2 \mathrm{D}}{\partial \mathrm{x}^2} \quad (4)$$

$$\hat{\ } \mathrm{x} = -\frac{\partial^2 \mathrm{D}}{\partial \mathrm{x}^2} \frac{\partial \mathrm{D}}{\partial \mathrm{x}} \quad (5)$$

From Equation 4 and 5, we will obtain equation 6

$$D\big(\hat{\ }(x)\big) = \mathrm{D} + \frac{1}{2} * \frac{\partial \mathrm{D}^{\mathrm{T}}}{\partial \mathrm{x}} * \hat{\ }\mathrm{x} \quad (6)$$

Direction task: the inclination size m(x, y) and direction θ(x, y) for each example point was resolved in this progression. The angle histogram (36 canisters) was built from slope orientation, weighted by Gaussian window (σ = 1.5).

(i) DESCRIPTOR:

The angle greatness and direction were processed around each example point in the area of key point. Figure 3 outlines a 2x2 descriptor cluster with 8 direction receptacles however the examinations were finished utilizing a 4x4 descriptor. This implies that we have a 4x4x8=128 component vector altogether. This vector was standardized to unit length to evade the effect of light changes. By determining the estimation of the unit highlight vector to be no bigger than 0.2, we will dispose of the impact of huge slope greatness.

Coordinating key points of inquiry picture with others in the information base was performed through utilizing closest neighbor edge with least Euclidean distance. Be that as it may, because of the foundation mess, numerous highlights couldn't discover any match, so another method was utilized called closest neighbor distance proportion (NNDR). It is a directed learning strategy, where preparing information is utilized for characterization.[3]

$$\mathrm{NNDR} = \frac{d1}{d2} = \frac{|D_A - D_B|}{|D_A - D_C|} \quad (7)$$

Where d1 and d2 are the closest and second closest neighbor distances. DA is the objective descriptor. DB and DC are its nearest two neighbors. shows the contrast between fixed limits, closest neighbor and NNDR.

3. K-MEANS CLUSTERING

K-means clustering is an illustration of an unaided AI that includes grouping the datasets without information on the marked class. K-implies grouping is an iterative, information apportioning calculation that relegates several perceptions to one of the k bunches characterized by centroids. In MATLAB, the centroids are dictated by utilizing a heuristic. [10] The calculation continues as follows:

1) Choose k introductory group habitats (centroid).

2) Compute highlight group centroid distances, all things considered, to every centroid.

3) There are two approaches (determined by Online Phase):

a) Batch update — Assign every perception to the group with the nearest centroid.

b) Online update — Individually allot perceptions to an alternate centroid if the reassignment diminishes the amount of the inside bunch, amount of-squares point-to cluster-centroid distances.

4) Compute the normal of the perceptions in each group to acquire the new number of centroid areas(k).

5) Repeat stages 2 through 4 until bunch tasks don't change, or the most extreme number of cycles is reached.

FUZZY LOGIC CLASSIFICATION ALGORITHM

This point of this cycle is to discover the kind of illness on strawberry leaves which are iron inadequacy or parasitic contamination. This calculation has five information sources. Two of the data sources allude to the iron inadequacy; The first is the number of green pixels of the picture and the subsequent information is the quantity of coordinated pixel between the red pixel of the picture and the green pixel in the wake of filling the openings on it. The leftover three sources of info allude to parasitic contamination the first is the quantity of green pixel of picture and the subsequent information is the zone of openings of the leaves and the third information is the coordinated pixels between the red pixels and the openings of the leaves. A Fluffy rationale-based calculation will yield two qualities which are numbers that signify the sort of illnesses of that strawberry plant on the off chance that it is sickness tainted. The fluffy rationale is made to improve arrangement productivity as far as precision and time, the estimated consequences of fluffy calculation are like the human vision which is superior to PC vision. Settling on choice for iron lack on plant leaves relies upon two factors.[5]

Approximate result for iron deficiency (ARID) = $\frac{ID(2)}{ID(1)}$ (8)

ID (1) is a variable that depends on input (1) which is the number of a green pixels of image.

$$ID(1) = I_1 \times input(1) \quad (9)$$

ID (2) is a variable that depends on input (2) which is the number of matched pixel between red pixel of image and the green pixel after filling the holes on it.

$$ID(2) = I_2 \times input(2) \quad (10)$$

where, I_1 and I_2 are the tuning parameters of fuzzy logic algorithm. $I_2 > I_1$

Making decision on fungal infection of infected leaves depends on three variables .

Approximate result for the fungal infection(ARFI) =$\frac{F1(2)+F2(3)}{F1(1)}$ (11)

FI (1) is a variable that depends on input (3) which is the number of a green pixels of image.

$$FI(1) = F_1 \times input(3) \quad (12)$$

FI (2) is a variable that rely on input (4) which is the area of holes of the leaves.

$$FI(2) = F_2 \times input(4) \quad (13)$$

FI (3) is a variable that depends on input (5) which implies the pixels matched between the red pixels and the holes of the leaves.

$$FI(3) = F_3 \times input(5) \quad (14)$$

where, F1, F2 and F3 are the tuning parameters of fuzzy logic.

$$F2 + F3 > F1 \quad (15)$$

One may notice that the abovementioned parameters were tuned subsequently over the optimal results.

4. DATASET DESCRIPTION

The dataset used is derived from the Plant village website. The Dataset contains plants disease images which are in the size of 512x512 and the extension is jpg. The dataset information is given below. This dataset used is derived from an online source and pre-processing mechanism is implemented. Pre-processing is used to extract the noise from within the image.

Fig 2 :COMMON RUST, LATE BLIGHT, CEDAR APPLE RUST

5. RESULTS:

Sensitivity (True Positive rate) quantifies the extent of positives that are recognized (for example the extent of the individuals who have some condition (influenced) who are effectively distinguished as having the condition).

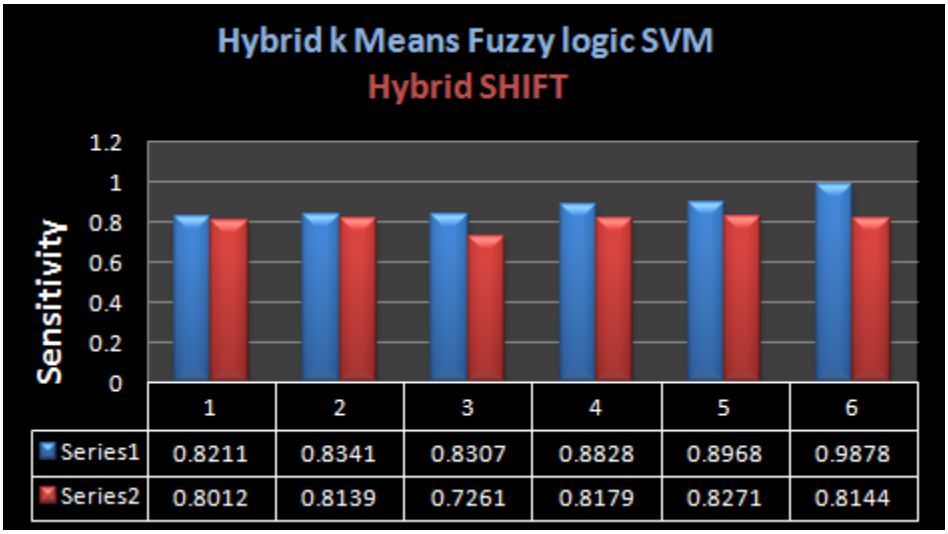

Fig 3 : SENSITIVITY OF TWO HYBRID ALGORITHM

Sensitivity=[a/(a+c)]×100

Specificity=[d/(b+d)]×100

Positive predictive value(PPV)=[a/(a+b)]×100

Negative predictive value(NPV)=[d/(c+d)]×100

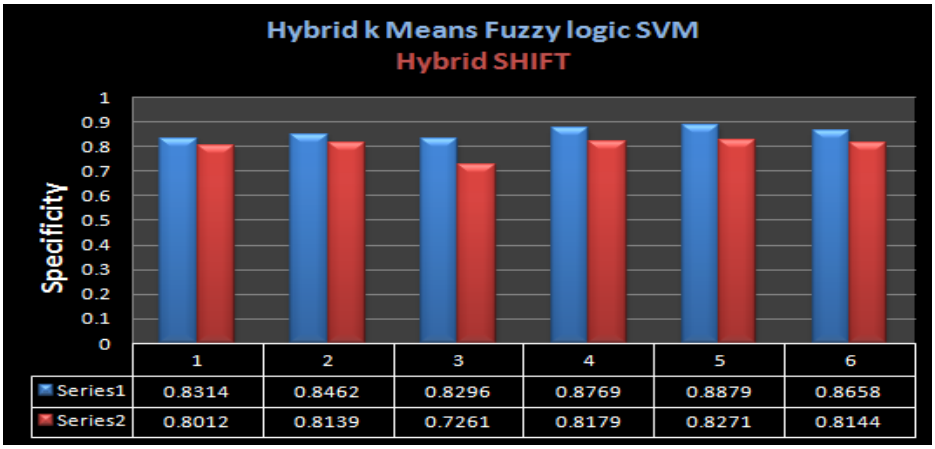

Fig4: SPECIFICITY OF TWO HYBRID ALGORITHM

Specificity (True Negative Rate) quantifies the extent of negatives that are accurately distinguished (for example the extent of the individuals who don't have the condition (unaffected) who are effectively recognized as not having the condition).

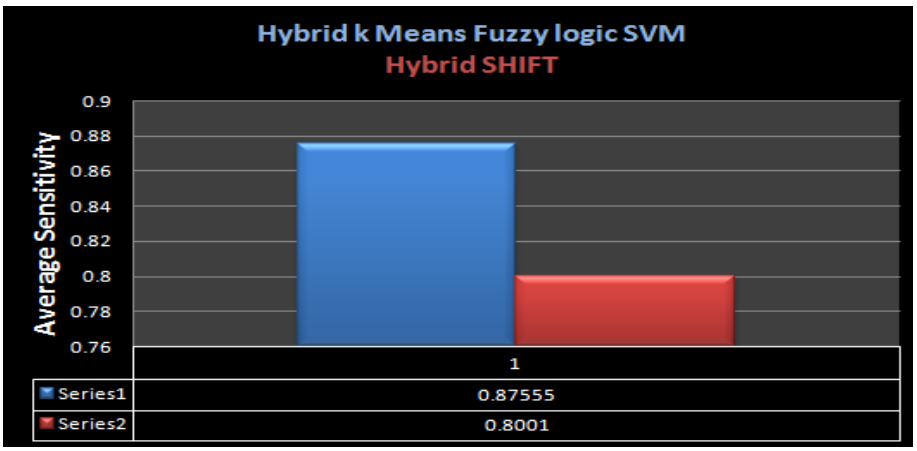

Fig5: AVERAGE SENSITIVITY

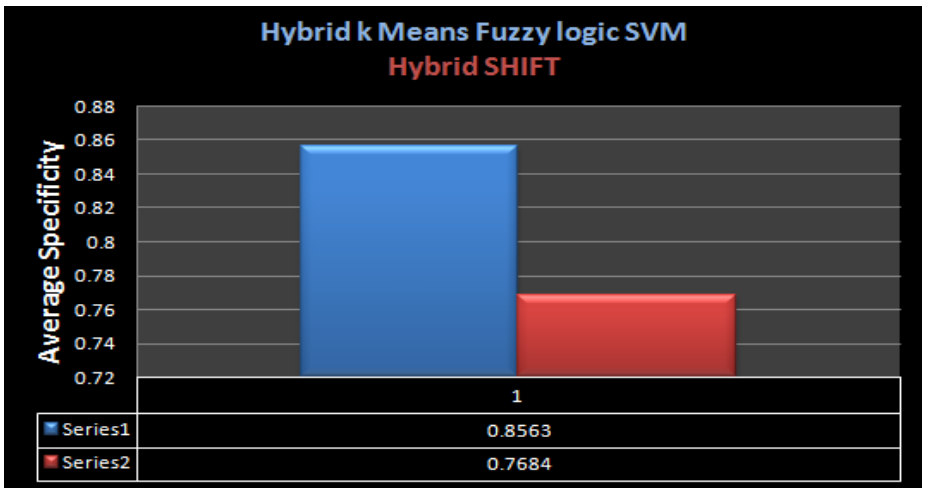

Fig6: AVERAGE SPECIFICITY

The above graph result shows the comparison of two hybrid algorithms (i.e. Hybrid K Means fuzzy logic SVM and Hybrid SHIFT) which mean of sensitivity and specificity with six diseases common rust, late blight, cedar apple rust, leaf spot early blight, and the above result suggest that HKMFLSVM gives good result for all disease.

6. CONCLUSION:

In general, the algorithm used can correctly distinguish between disease part of a plant, the area affected by six categories as well as the background. A high percentage obtained for both classification of background images and the leaf part of the plant. The proposed system for the examination of plant disease infection was investigated and the identification of disease-infected leaves at a different field and the present research work overcoming the challenges for on-site disease detection will result in crop health monitoring in plants leaf. In this paper, the efforts were put into proposing a method to be deployed on the hybrid algorithm and the proposed algorithm HKMFLSVM achieved a good result.

REFERENCES

1. Y. Tao, M. Skubic, T. Han, Y. Xia, and X. Chi, Evaluation of SIFT Based Descriptors for Object Recognition,IMECS, vol. II, no. 2, pp. 17-20, 2010.
2. Khan, N.Y.; McCane, B.; Wyvill, G, "SIFT and SURF performance evaluation against various image deformations on benchmark dataset", Proceedings of the 2011 International Conference on Digital Image Computing Techniques and Applications (DICTA), Queensland, Australia, 68, December 2011; pp. 501-506.
3. P M Panchal, S R Panchal, S K Shah, A Comparison of SIFT and SURF International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 2, pp. 323-327, April 2013.
4. Bai, X., Li, X., Fu, Z., Lv, X., Zhang, L., 2017. A fuzzy Clustering Segmentation Method Based on Neighbourhood Grayscale Information for Defining Cucumber Leaf Spot Disease Images. Computers and Electronics in Agriculture 136, 157-165
5. Barbedo, A., 2016. A review on the Main Challenges in Automatic Plant Disease Identification Based on Visible Range Images. Biosystems Engineering 40, 52–60.
6. Boissard, P., Martin, V., Moisan, S., 2008. A cognitive Vision Approach to Early Pest Detection in Greenhouse Crops. Computers and Electronics in Agriculture 62, 81–93
7. Tripathi, M. K., Maktedar, D. D.: Recent machine learning based approaches for disease detection and classification of agricultura
8. products,International Conference on Computing Communication Control and automation (ICCUBEA), pp. 1--6, (2016).
9. Khirade, S. D., Patil, A. B.: Plant Disease Detection Using Image Processing, International Conference on Computing Communication Control and Automation, pp. 768--771 (2015).
10. Ivey, M. L., Nava-Diaz, C., Miller, S.: Identification and Management of Colletotrichum acutatum on Immature Bell Peppers, Plant Disease, 88(11) pp. 1198-1204 (2004).
11. Russell, S., norvig, P.: Aritifical Intelligence, A Modern Approach, 3rd ed., New Jersey: Pearson (2010).
12. Sklansky, J.:Image Segmentation and Feature Extraction, In IEEE Transactions on Systems, Man, and Cybernetics, vol. 8, no. 4, pp.237--247, (1978).
13. Padol, P. B., Yadav, A. A.: SVM classifier based grape leaf disease detection, Conference on Advances in Signal Processing (CASP), pp. 175--179 (2016). 61